

Ch. 16: Random Variables

* A numeric value based on the outcome of an event


* We use capital letters to denote random variables (like X, Y, Z)

* Example: tossing a coin 3 times and recording the number of heads: $X = 0, 1, 2, 3$ $X=2.5$

* Example: # of children in a family: $X = 0, 1, 2, 3, \dots$ $X=2.5$

Random Variables: 2 types

	Discrete R.V. *	Continuous R.V. ^{Unit 1}
What is it?	* Can list all outcomes	* Outcomes are in an interval * cannot list all
Examples:	* # of heads in 3 coin tosses * # of children in a family	* # of hours a light bulb lasts $(0, 2000)$ * Time to run a mile $(3:43, \infty)$ $(4:13,)$

	DISCRETE	CONTINUOUS												
Distribution/ Function	<ul style="list-style-type: none">* Probability Model <table border="1"><tr><td>X</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>P(X)</td><td>0.1</td><td>0.3</td><td>0.5</td><td>0.03</td><td>0.07</td></tr></table> <ul style="list-style-type: none">* Probability histogram <p>$P(X=2) = 0.30$</p>	X	1	2	3	4	5	P(X)	0.1	0.3	0.5	0.03	0.07	<ul style="list-style-type: none">* Continuous Distrib.  <ul style="list-style-type: none">* Ex: <u>normal distrib.</u>, <u>uniform distrib.</u>
X	1	2	3	4	5									
P(X)	0.1	0.3	0.5	0.03	0.07									
Properties of Distrib.	<ul style="list-style-type: none">* Sum up to 1 <p>$P(X < 4) = 0.90$</p> <ul style="list-style-type: none">* All <u>disjoint prob.</u> <p>Add prob. to get sums</p>	<ul style="list-style-type: none">* Area under curve = 1 <p>$N(100, 5)$</p> <ul style="list-style-type: none">* No individual probabilities!! <p>$P(X \geq 10) =$</p>												

We will focus on DISCRETE Random Variables in this chapter



ample:

- 1) Let's play a game! You pay \$5 to play. In the game, you get to draw one card from the deck.
- If you draw the ace of hearts, you win \$100
 - If you draw any other ace, you win \$10
 - If you draw any other heart, you win \$5
 - Any other card, you win nothing
 - We are interested in the amount that we GAIN

a) Create the probability distribution below:

X = values of the variable. The outcomes
P(X) = the probabilities of the values. Must sum to 1.

(X)	\$95	\$5	\$0	-\$5
P(X)	1/52	3/52	12/52	36/52

= 1

- b) What is the chance that you go home with at least some money? $P(X > 0) = 4/52$
- c) What is the chance that you ^{GAIN} at least \$10? $P(X \geq 10) = 4/52$
- d) What is the probability of not gaining any money? $P(X \leq 0) = 48/52$
- $P(X=95)$ $P(X=95)$ $P(X \geq 5)$ $P(X=5 \cup 95)$

Mean/Std. Deviation

MEAN =

* Notation:

Expected Value = the value we would expect to get if we played the game many, many, many times. Theoretical long-run average.

Formula (how to find it):

- X is a Discrete R.V. with the following distr.:

X	X1	X2	X3 ...	Xn
P(X)	P(X1)	P(X2)	P(X3) ...	P(Xn)

- Find the mean (expected value) by doing the following:

CALCULATOR:

- L1 = values of variable (top row of table)
- L2 = Probabilities of each value (bottom row of table)
- STAT \rightarrow CALC \rightarrow 1-Var-Stats L1, L2
- Mean = \bar{x} Std. Deviation = σ

ample:

X	0	1	2	3	4	5
P(X)	0.05	0.12	0.18	0.2	0.4	0.05

What is the mean (expected value) and std. deviation?

ample: Find the expected value (mean) and the std. deviation of the first game above.

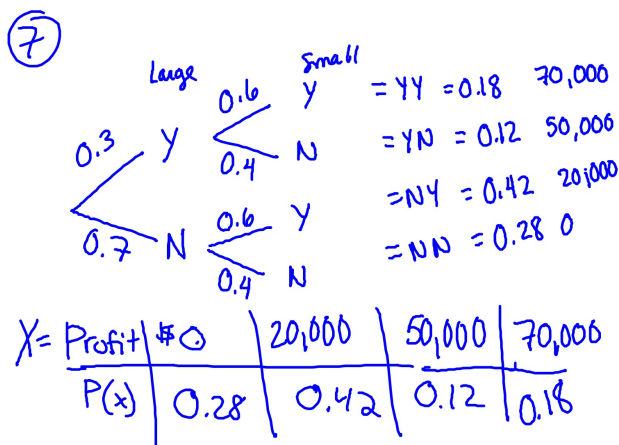
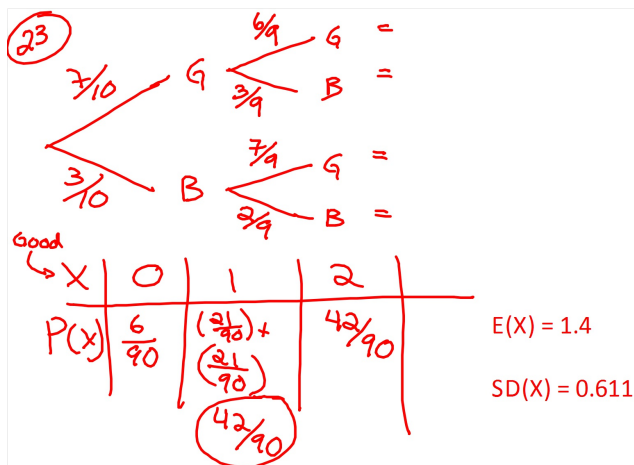
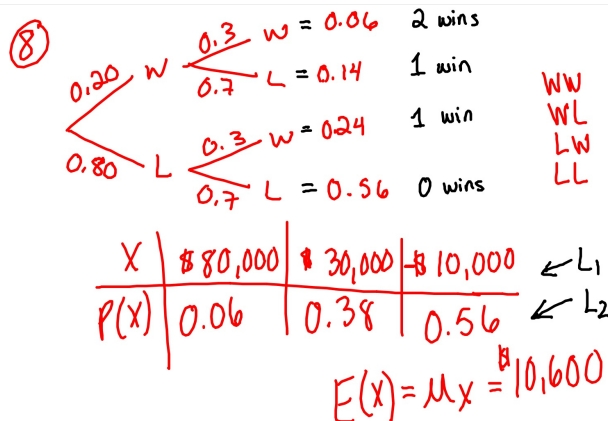
ample: New game! We are rolling a die. We are again interested in the amount ~~\$\$\$~~ it costs \$5 to play

GAIN

- Roll a 6, you win \$10
- Roll a 5, you win \$7
- Roll a 3 or 4, you win \$5
- Roll a 1 or 2, you win nothing

ite the prob. distribution and find the expected value (mean) and standard deviation of the game:

Book examples: p. 383 #5, 8, 23, 7



Transforming Random Variables

- We may want to alter a random variable (add, subtract, multiply, divide the variable)
- We may want to combine it with another random variable

Let X and Y be independent random variables

Let "a" and "b" be fixed numbers

LESS

Multiplication and addition of a variable:

$$\mu_X = 10$$



MEAN: $\mu_{a+bX} = a + b \cdot \mu_X$

VARIANCE: $\sigma_{a+bX}^2 = b^2 \sigma_X^2$ $\sigma_{a+bX} = b \cdot \sigma_X$

$$\sqrt{b^2 \sigma_X^2}$$

$$b \sigma_X$$

REMEMBER:

Std. Deviation = $\sqrt{\text{variance}}$

$$\sigma = \sqrt{\sigma^2}$$

(X and Y)

Combining 2 or more variables together:

REMEMBER:

Std. Deviation = $\sqrt{\text{variance}}$

$$\sigma = \sqrt{\sigma^2}$$

MEAN: $\mu_{X+Y} = \mu_X + \mu_Y$

$$\sigma_X = 3 \quad 9$$

$$\sigma_Y = 4 \quad 16$$

$$\mu_{X-Y} = \mu_X - \mu_Y$$

VARIANCE:

$$\sigma_{X+Y}^2 \rightarrow \sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

$$\neq \sigma_X + \sigma_Y$$

$$\sigma_{X-Y}$$

EXAMPLES:

- 1) Random variable X has a mean of 6.2 and a standard deviation of 3.1.
- a. Find the new mean and standard deviation if we multiply by 3
- $$\mu_{3X} = 3 \cdot 6.2 = 18.6$$
- $$\sigma_{3X} = 3 \cdot 3.1 = 9.3$$
- b. Find the new mean and standard deviation if we subtract 10
- $$\mu_{X-10} = 6.2 - 10 = -3.8$$
- $$\sigma_{X-10} = 3.1$$
- c. Find the new mean and standard deviation if we multiply by 5 and add 10.
- $$\mu_{5X+10} = 41$$
- $$\sigma_{5X+10} = \sigma_{5X} = 15.5$$

- 2) Random variable Y has a mean of 3.4 and a standard deviation of 1.4.

- a. Find the mean and standard deviation of $X+Y$
- $$\mu_{X+Y} = 9.6$$
- $$\sigma_{X+Y}^2 = (3.1)^2 + (1.4)^2 = 11.57 = 3.401$$
- $$\sigma_{X+Y} = 3.401$$
- b. Find the mean and standard deviation of $X-Y$
- $$\mu_{X-Y} = 6.2 - 3.4 = 2.8$$
- $$\sigma_{X-Y} = 3.401$$
- c. Find the mean and standard deviation of $2X+3Y$
- $$\mu_{2X+3Y} = 2(6.2) + 3(3.4) = 22.6$$
- $$\sigma_{2X+3Y}^2 = (2 \cdot 3.1)^2 + (3 \cdot 1.4)^2 = 56.08 = 7.489$$
- $$\sigma_{2X+3Y} = 7.489$$
- d. Find the mean and standard deviation of $3X+Y-4$
- $$\sigma_{3X+Y-4} = \sqrt{(3 \cdot 3.1)^2 + 1.4^2}$$

New game! A single dice is rolled and the following occurs:

- Roll a 6, get 40 points
- Roll a 4 or 5, get 10 points
- Roll a 1, 2, or 3, get 0 points

Write the probability distribution

pts.	40	10	0
P(p)	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$

Find the mean and standard deviation of the game

$$\mu_P = 10 \text{ pts.} \quad \sigma_P = 14.142 \text{ pts.}$$

Suppose the points are doubled. Find the new mean and standard deviation

$$\mu_{2P} = 20 \text{ pts.} \quad \sigma_{2P} = 28.284 \text{ pts.}$$

Supposed the game is played twice (independently). What is the mean and standard deviation of this?

$$\mu_{P+P} = 20 \text{ pts.}$$

$$\sigma_{P+P} = \sqrt{14.142^2 + 14.142^2} = 20 \text{ pts.}$$

More examples:

- Suppose that X is a random variable with $\mu_X = 10$ and $\sigma_X = 2$. Find the following:
 - μ_{5X}
 - μ_{X-3}
 - μ_{5X-3}
 - σ_{5X}
 - σ_{X-3}
- Suppose that X is the random variable from above, and that Y is independent from X with $\mu_Y = 15$ and $\sigma_Y = 3$. Find the following:
 - μ_{X+Y}
 - μ_{X-3Y}
 - σ_{X+Y}
 - σ_{X-3Y}
 - σ_{5X-2Y}
- Suppose the mean SAT verbal score is 425 with standard deviation 100, while the mean SAT math score is 475 with standard deviation 100. What can be said about the mean and standard deviation of the combined math and verbal score?

Book examples: p. 384 #16 & 32, 24, 25, 38, 40

16) (a) $\mu_X = 2.25$ red lights
 (b) $\sigma_X = 1.26$ red lights

$\mu_{X+X+X+X+X} = 11.25$ red lights
 $\sigma_{X+X+X+X+X} = \sqrt{5(1.26)^2}$

32) $\mu_{X+X+X+X+X} = 11.25$ red lights
 $\sigma_{X+X+X+X+X} = 2.82$ red lights

38) $\mu_D = \$100$

$\sigma_D = \$30$

$\mu_C = \$120$

$\sigma_C = \$35$

(a) $\mu_{D-C} = -\$20$

(b) $\sigma_{D-C} = \sqrt{30^2 + 35^2}$
 $= \$46.10$

(c) Normal $N(-20, 46.10)$

$P(D > C)$
 $-C$

$P(D-C > 0) = \text{normcdf}(0, \infty, -20, 46.10)$

25) (a) $\mu = 30$ $\sigma = 6$ (b) $\mu = 26$ $\sigma = 5$ (c) $\mu = 30$ $\sigma = 5.39$ (d) $\mu = -10$ $\sigma = 5.39$

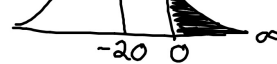
(e) $\mu = 20$ $\sigma = 2.83$

38) $\mu_D = \$100$ $\mu_C = \$120$
 $\sigma_D = \$30$ $\sigma_C = \$35$

(a) $\mu_{D-C} = -\$20$

(b) $\sigma_{D-C} = 46.098 \rightarrow \46.10

(c) $\text{normcdf}(0, \infty, -20, 46.10)$
 33.22%



40) a) $D + D + C = \text{total cost}$

b) $\mu_{D+D+C} = 100 + 100 + 120 = \320

$\sigma_{D+D+C} = \sqrt{30^2 + 30^2 + 35^2} = \55

$N(320, 55)$

c) $P(\text{total} > 400) =$

$= \text{normcdf}(400, \infty, 320, 55)$

$= 0.0729$